DOCUMENT RESUME

ED 431 028 TM 029 861

AUTHOR Ang, Rebecca P.; Woldbeck, Tanya F.; Hughes, Jan N.
TITLE Impact of Modality on Skills Training for Youth with Externalizing Problems: A Meta-Analysis.
PUB DATE 1999-04-00
NOTE 15p.; Paper presented at the Biennial Meeting of the Society for Research in Child Development (Albuquerque, NM, 15-18, April 1999).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Adolescents; *Antisocial Behavior; Behavior Problems; Elementary Secondary Education; Emotional Problems; Meta Analysis; *Outcomes of Treatment; Program Effectiveness; Program Evaluation; *Skill Development; *Training; Urban Youth; Young Adults

ABSTRACT

A meta-analysis (37 published studies) that summarized treatment outcomes associated with skills training with antisocial youth was performed. As is consistent with the hypothesis, results indicate that skills training interventions delivered in the context of homogeneous groups of deviant peers produced smaller benefits than did skills training interventions delivered in the context of mixed groups of prosocial and deviant peer, or individual treatment. Also, as expected, treatment provided in the context of deviant-only groups attenuated treatment benefits more for more severely disordered groups such as those who are incarcerated or placed in a class for behavioral or emotional problems, than for youth who might only be at-risk for such conditions. (Contains 1 table and 50 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

RUNNING HEAD: IMPACT OF MODALITY ON SKILLS TRAINING

ED 431 028

Impact of Modality on Skills Training for Youth with Externalizing Problems: A Meta-Analysis

Rebecca P. Ang, Tanya F. Woldbeck, and Jan N. Hughes
Texas A&M University

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Rebecca Ang

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Poster presented at the biennial meeting of Society for Research in Child Development, Albuquerque, NM, April 1999.

TM029861

BEST COPY AVAILABLE

Abstract

A meta-analysis (37 published studies) which summarized treatment outcomes associated with skills training with antisocial youth was performed. Consistent with our hypothesis, results indicated that skills training interventions delivered in the context of homogeneous groups of deviant peers produced smaller benefits than did skills training interventions delivered in the context of mixed groups of prosocial and deviant peers, or individual treatment. Also, as expected, treatment provided in the context of deviant-only groups attenuated treatment benefits more for more severely disordered youth such as those who are incarcerated or placed in a class for behavioral or emotional problems, than for youth who might only be at-risk for such conditions.

Impact of Modality on Skills Training for Youth with Externalizing Problems: A Meta-Analysis

Research on interventions designed to reduce childhood aggression and prevent adolescent delinquency and substance abuse has expanded rapidly in the past decade. One of the most popular intervention methods in schools and clinics is group-based skills training in which a group of youth with conduct problems are taught a set of social or problem-solving skills to help them better negotiate problem situations without using aggressive means (Kazdin, 1997). Group-based skills training is popular in part because of its ease and efficiency of administration. Whereas some group-based interventions have documented positive outcomes, others have failed to do so, and some interventions result in detrimental outcomes (see Arnold & Hughes, in press).

In a meta-analytic investigation of group social skills training with children experiencing a range of problems, Beelmann, Pfinsten, and Losel (1994) found that group-based skills training resulted in modest short-term gains but that long-term benefits were generally lacking. Also, short-term gains varied as a function of assessment method, with larger gains for measures of targeted skills and smaller gains on measures of socially consequential outcomes. Because the Beelmann et al. meta-analysis included children with problems other than externalizing disorders, their findings may be more or less characteristic of skills training with antisocial[1] youth. A quantitative summary of the effects of skills training with deviant youth would assist in evaluating the efficacy of this intervention modality with this population. In particular, it is important to determine if intervention effectiveness differs based on characteristics of participating youth and the training program.

In a narrative review of group-based skills training with aggressive youth, Arnold and Hughes (in press) argue that whether or not skills training occurs within the context of deviant-only groups may affect its benefits. Specifically, they suggest that the expected benefits of skills training with this population may be diluted as a result of unintended negative effects of aggregating deviant youth for purposes of skills training. Arnold and Hughes (in press) underscored the necessity of systematic research on the effect of grouping deviant peers for skills training interventions in order to test the hypothesis that aggregating delinquent youth attenuates treatment gains. They recommended a meta-analytic investigation of the possible moderating role of group composition (i.e., deviant-only treatment versus non-aggregated treatment) on the effectiveness of skills training with aggressive youth. This article attempts to determine via a comprehensive review of all controlled outcome studies to date of social skills interventions with antisocial youth, if positive treatment effects are greater for skills training interventions that provide individualized treatment or treatment in groups comprised of both prosocial and aggressive peers, than are treatment effects for skills training in interventions that provide treatment in homogeneous groups of antisocial children. We were also interested in investigating whether selected client characteristics moderated the effectiveness of skills training and whether client characteristics exerted such an influence differently based on group modality. Specifically, we expected that deviant-only group treatment would be less successful with youth whose antisocial behaviors were less severe, versus youth with more severe antisocial behavior. This expectation was based on a finding that moderately aggressive boys are most susceptible to the deleterious influence of aggressive friends (Vitaro, Tremblay, Kerr, Pagani, & Bukowski, 1997). We also expected that age might be a moderating variable such that treatment modality would not account for differential treatment benefit for older

---

[1] Throughout this manuscript we use the terms "antisocial" and "deviant" to refer to youth with a range of externalizing behavior problems including aggression, defiance, stealing, and lying; when discussing specific studies, we use terms that characterize the subjects in the particular studies.

youth (age 13-18). We expected that by mid-adolescence, affiliation patterns and identification with conforming or non-conforming peer groups would be more stable; consequently, mixing procial peers and antisocial peers for treatment would have little impact on youth's peer networks or attitudes toward conventional versus antisocial behavior.

With the development of meta-analytic techniques, it is now possible to integrate findings across multiple studies and to systematically compare findings across dimensions such as outcome type, for example. The basis for analysis is the effect size, which is an estimate of the magnitude of the treatment effect adjusted for sample variability. The aggregation of results from different studies is a major advantage of meta-analysis over the traditional narrative literature review. A more direct statistical comparison of studies can be made with more control over possible bias inherent in the narrative review. Schmidt (1992, 1996) pointed to another advantage of a meta-analytic review: statistical significance testing for interpreting the data no longer plays such a dominant role. Statistical significance testing can lead researchers to mistakenly conclude that no relationship exists between two variables of interest (Type II error). One limitation of the meta-analytic approach however, is that many studies do not provide sufficient information or information of sufficient detail to permit inclusion in a review (Lipsey & Wilson, 1993).

Method

Selection of Studies

This review was restricted to published studies based on an assumption that published research undergo a review process that controls for the quality of research studies. We defined social skills training as behavioral and/or cognitive interventions that were explicitly directed toward training or modifying cognitive (e.g., problem-solving skills) and/or affective (e.g., anger control) components of social behavior. Other criteria for inclusion of a study in the review were as follows:

1. Selection was based on mean age of total sample or reported school grade. Subjects were between 6-to-18 years of age, or were in school grades ranging from 1 through 12.
2. Subjects with described as having externalizing behavior problems including (a) childhood aggression, (b) conduct disorder, (c) oppositional defiant disorder, (d) antisocial behavior, (e) violent behavior, or (f) adolescent delinquency. Studies with subjects described as hyperactive or experiencing peer rejection were excluded in this review unless the subjects also presented with conduct problems.
3. Studies used an experimental or quasi-experimental design with at least one control group. Studies that included group comparisons with only a nondeviant control group, and single group designs were excluded.
4. Studies involved group or individual treatment. Group composition could either be deviant-only peers, or mixed groups of deviant and prosocial peers.
5. Outcome assessment reported quantifiable measures of social or behavioral adjustment.
6. Studies were published between mid-1970s through 1997 in English.

Literature Search Procedure

With the selection criteria listed in the preceding section, a total of 37 studies (5 individual, 5 mixed-group, and 27 deviant-group) were identified (marked with an asterisk in the Appendix). The studies were identified through PsycINFO and ERIC databases which index academic and professional literature in psychology and education, as well as related disciplines such as psychiatry, medicine, nursing, and sociology. Using a computer search, the following keywords were used in various combinations: social skills training, problem solving skills training, cognitive therapy, behavioral therapy, cognitive behavioral therapy, group therapy, group intervention, conduct

disorder, aggression, delinquency, antisocial, violence, deviant, and prosocial. In addition, the references from these identified studies were inspected to locate studies appearing in other publications.

Coding Procedure

Each study was coded with respect to subject, treatment, and outcome measure characteristics. Most classifications were straightforward but a few warrant explanation. We wanted to assess how treatment modality and type of population relate to treatment outcome. Treatment modality was coded as deviant-only group treatment, individual treatment, or mixed (deviant and prosocial) group treatment. In order to assess the magnitude of treatment effects associated with homogeneous group treatment of deviant children versus either individual treatment or treatment of children in mixed deviant and prosocial groups, we collapsed the individual treatment and mixed group treatment into a single classification (non-aggregated treatment) prior to conducting our analyses. Collapsing across studies evaluating individual and mixed group treatments was also necessary in order to have reasonable power to detect differences based on treatment modality. Population type was classified as either preventive or clinical. The preventive population included children and youth that manifested aggressive, disruptive and delinquent behaviors which have not been clinically diagnosed, and who are not incarcerated or institutionalized. The clinical population included aggressive children and youth that have either been diagnosed as conduct disordered or are juvenile offenders that have been incarcerated or institutionalized in a psychiatric treatment facility. Type of outcome measures was classified into five categories: (a) behavior rating scales (e.g., teacher or parent rating scales), (b) behavior observations (e.g., time-sampled ratings of children's behavior and role-play performance) (c) self-report (e.g., measures assessing self-esteem) (d) problem-solving skills (e.g., measures assessing children's problem-solving skills via the presentation of hypothetical situations or vignettes), and (e) socially consequential measures (e.g., sociometrics, recidivism). Several methods by which children were referred included being in a special program (e.g., The Think First Program), the use of rating scales, behavioral indicators, teacher nomination, teacher nomination plus the use of rating scales, and teacher and peer ratings.

Interrater Agreement

The first author performed the coding for all studies. Eighteen studies (48.6%) were randomly selected and coded by another author to test inter-rater agreement. There was perfect agreement for treatment modality ($\kappa = 1.00$) and type of population ($\kappa = 1.00$). The two raters achieved kappas of .79 for type of outcome measures and .90 for method of referral. Disagreements in coding between the raters were resolved through discussion.

Estimating Treatment Effects

Effect sizes were estimated using procedures suggested by Glass, McGaw, and Smith (1981). In each calculation, effect size was computed as treatment group mean minus control group mean divided by the control group standard deviation. We used the control group standard deviation instead of the pooled within-group standard deviation as the denominator because we agree with Bergin and Lambert (1978) that one consequence of therapy is an increase in behavioral variability. Thus, use of the pooled within-group standard deviation may cause statistical and interpretational problems (Smith, Glass, & Miller, 1980) which we attempted to avoid.

For some outcome measures, higher numbers indicated greater improvement but for other measures lower numbers indicated greater improvement. Effect sizes were calculated in a consistent

manner such that positive scores indicate that the treatment group improved more than the control group and negative scores indicate that the control group improved more than the treatment group.

Most effect sizes were calculated from means and standard deviations or raw data reported in the study. When this information was unavailable, effect size was estimated from other reported statistics (e.g., t, F, or chi-square). Various techniques for estimating such effect sizes are reported in previous work (see Glass et al., 1981, chap. 5).

In some studies, findings for all outcome measures that had been used were either not reported or reported as nonsignificant. In such cases, an exact effect size cannot be computed. However, not including these outcome measures in the analyses would have artificially inflated the overall estimate of effect size because researchers were more likely to provide complete or adequate information on those measures that demonstrated statistically significant or large treatment effects. Thus, when the results from an outcome measure were not reported or were reported only as nonsignificant, we conservatively estimated the effect size to be zero.

Most studies compared treatments on more than one outcome measure. Because multiple effect sizes derived from the same study may not represent statistically independent observations, an analysis based on such nonindependent observations can underestimate error variance and inflate tests of statistical significance (see Glass et al., 1981, chap. 6). To avoid this problem of nonindependence of observations, we averaged multiple effect sizes obtained from individual measures within the same treatment comparison. However, separate means for different types of outcome measures were also calculated because we wished to assess treatment effects associated with particular types of outcome measures. We calculated separate means for five types of outcome measures: behavior ratings, behavior observations, self-report, problem-solving skills, and socially consequential measures. Multiple effect sizes obtained for each outcome measure type (e.g., behavior ratings) within a study were averaged to obtain a single effect size for that outcome measure type. For example, if behavior ratings yielded three effect sizes within a study, these effect sizes would be averaged to yield a single effect size value associated with behavior ratings for that study.

Several studies reported the results of more than one treatment comparison. This problem of nonindependence was treated in a manner similar to that used for multiple outcome measures. For example, if a study yielded two different comparisons of treatment (e.g., social skills training and social skills training with in vivo practice) to no treatment, we averaged the results of these comparisons to get a single value for the treatment group to be used in the analyses.

Another issue we encountered was the use of multiple control groups. Some studies used a single control group whereas others utilized two types of control groups (e.g., attention-placebo, no treatment control). In studies where more than one control group was used, we used the more stringent control group as our control comparison against the treatment comparison. For example, if a study had three groups, social skills training group, attention-control, and no treatment control, the more stringent attention-control group would serve as our control comparison.

In a typical meta-analysis, an index of effect size is used to summarize the results of each study, and effect size indices may then be averaged to obtain an overall estimate of effect magnitude. Conventional statistical methods such as analysis of variance (ANOVA) and multiple regression are then used to study the variation in the effects across studies. Hedges and Becker (1986) argued that conventional analyses frequently involve serious violations of the assumptions of these techniques, and have demonstrated that statistical methods they developed overcome both conceptual and statistical problems posed by conventional statistical analyses of effect sizes. Conceptually, conventional analysis lacks the ability to test the consistency of effect sizes across studies. This limitation is important because combining effect sizes across studies makes sense only if the studies have a common population effect size. Consequently, it is impossible to construct a

test for whether the systematic variation in the effect sizes is larger than the nonsystematic variation exhibited by those effect sizes. Conventional methods are also problematic for statistical reasons. Conventional statistical procedures (e.g., ANOVA, multiple regression analysis) rely on parametric assumptions about the data that are not fulfilled for effect size data. These procedures require that the unsystematic variance associated with every observation is the same. The unsystematic variance of estimates of effect size is proportional to 1/n, where n is the sample size of the study on which the estimate is based. Therefore, if the sample sizes of the studies vary widely, which is usually the case, the effect size estimates will have different error variances. The F-test is not necessarily robust to severe violations of the homogeneity-of-variance assumption in ANOVA and regression analyses.

Hedges and Becker (1986) developed new statistical procedures to overcome the conceptual and statistical problems associated with conventional analytic procedures. Hedges' (1981) correction factor is applied to the effect size estimate obtained via the procedure outlined by Glass et al. (1981) to yield the unbiased estimator d. The variance of d is completely determined by sample sizes and the value of d. Therefore, it is possible to determine the sampling variance of d from a single observation. This ability to determine the non-systematic variance of d from a single observation of d is the crux of modern statistical methods for meta-analysis. Refer to Hedges and Becker (1986) for a comprehensive review and demonstration of these techniques. The authors used Hedges and Becker's techniques in the computation of the unbiased estimator d and in all effect size analyses.

Results

Sample and Treatment Characteristics

Five studies were classified as individual treatment, five studies were classified as mixed group treatment, and 27 studies were classified as deviant-only treatment. For posttreatment data, 11 effect sizes were obtained from individual treatment studies, 14 effect sizes were obtained from mixed group treatment studies, and 72 effect sizes were obtained from deviant-only group treatment studies. The mean age of the subjects was 11.54 years (SD = 2.84; range = 6.0 to 18.1). On average, 85% of the youngsters sampled were male. Of the 37 studies, 15 did not report information regarding ethnicity. The remaining studies had ethnicity breakdowns with the following mean percentages: 47% Anglo-Americans, 51.9% African-Americans, 1.8% Hispanics, and 1.4% Others. Across studies, the mean number of sessions per week was 1.83 (SD = 1.03; range = 1 to 5) and the mean number of minutes per session was 59.16 (SD = 30.41; range = 20 to 180). The average treatment duration in weeks was 13.26 (SD = 17.09; range = 3 to 104). Eighteen studies (48.6%) provided information on follow-up treatment, and 19 studies (51.4%) did not have follow-up. For the 18 studies that had follow-up treatment, the mean length of time between posttreatment and follow-up in months was 5.18 (SD = 9.18; range = 1 to 36).

Therapist Training

Ten studies did not provide information (27.0%) on the experience level of the therapist. Of the 27 studies (73.0%) that did report such information, therapy was provided by professionals in 19 studies (50%), graduate students in 7 studies (18.4%), and university professors in 1 study (2.6%).

Testing Homogeneity of Effect Size

The unbiased estimator d as described previously, is based upon effect size estimates of independent samples. A weighted average D is a precise combination of values of d that takes into account the variances of d. Such a combination of effect sizes across studies makes sense only if the studies shared a common underlying population effect size. A test of the homogeneity of effect size

involves the computation of the Q_T statistic, which is the weighted sum of squares of effect sizes estimates about the weighted mean D. If all the studies share a common underlying population effect size, then Q_T has approximately a chi-square distribution with k-1 degrees of freedom. However, if sufficient heterogeneity exists, then Q_T will tend to be larger than expected by chance. Thus the test for the homogeneity of effect size will be rejected at the significance level α if Q_T exceeds the $100(1 - \alpha)$ percent critical value of the chi-square distribution with k-1 degrees of freedom. The overall effect size across the 37 studies was 0.49. Because the homogeneity test revealed that the effect sizes across the studies differed beyond chance ($Q_T = 72.90$, $p < .05$), any attempt to interpret the overall average effect size may be misleading, and hence an investigation of factors that may moderate effect size is warranted.

Effect of Treatment Modality on Outcome Measures

The authors sought to determine the effect of modality (individual and/or mixed versus deviant-only) on various outcome variables using the procedures outlined by Hedges and Becker (1986). This procedure is an analogue to ANOVA for effect sizes which permits the authors to test the significance of variation between groups of effect sizes, and to test if the remaining variation within groups of effect sizes is significant. The ANOVA for effect sizes involves partitioning the overall homogeneity statistic Q_T into the between-group homogeneity (Q_B) and the within-group homogeneity (Q_W). The between-group homogeneity statistic Q_B is analogous to the F statistic for testing for between-group differences in a conventional ANOVA. When there are p groups the statistic Q_B has approximately a chi-square distribution with p-1 degrees of freedom. When testing for between-group differences, Q_B is compared with the $100(1 - \alpha)$ percent critical value of the chi-square distribution with p-1 degrees of freedom, and if Q_B exceeds the critical value, the between-group difference is significant at level α. The within-group homogeneity statistic Q_W is the sum of the homogeneity statistics calculated for each of the p groups as if each group were an entire collection of studies.

There are two sub-categories in the non-aggregated treatment modality classification, individual treatment and mixed (deviant and prosocial) group treatment. The mean effect size for the five individual treatment studies (0.72) and the five mixed group treatment studies (0.67) did not differ from each other statistically, $t(20) = 0.32$, ns. Because the effect sizes for these two categories were comparable, they were collapsed into a single classification (non-aggregated treatment) for the analyses.

Modality (non-aggregated versus deviant-only) across all outcome variables was found to be statistically significant ($Q_B = 27.57$, $p < .05$). Table 1 reports effect sizes by treatment modality for all effects. The pattern of means was consistent with the hypothesis; the means were 0.42 and 0.69 in the deviant-only group and the non-aggregated group, respectively. We next tested whether the advantage of non-aggregated grouping was consistent across different types of outcome measures. Modality was found to be statistically significant for posttreatment outcome variable behavior observations ($Q_B = 17.69$, $p < .05$). The mean for the deviant-only group was 0.39 and the mean for the non-aggregated group was 0.80. The mean effect size for the deviant-only group (0.14) differed significantly from the mean effect size for the non-aggregated group (0.66) for outcome variable problem solving skills, ($Q_B = 17.87$, $p < .05$). Similarly, the effect of modality was found to be statistically significant for socially consequential outcome measures ($Q_B = 8.06$, $p < .05$). Once again, the pattern of means was in a direction consistent with that of the hypothesis; the means were 0.95 and 1.34 for deviant-only and non-aggregated group respectively. Differences between deviant-only modality and the non-aggregated modality for outcome variable categories behavior

ratings and self-report did not reach statistical significance at $\alpha = .05$ but were in the same direction as the overall findings.

In addition to examining the impact of modality on posttreatment effect sizes, we conducted similar analyses for 38 follow-up effect sizes which were obtained from the studies reporting follow-up data. The overall effect size for follow-up data was found to be 0.30. Since the test of homogeneity revealed that sufficient heterogeneity existed among the effect sizes ($Q_T = 55.81$, $p < .05$), the overall effect size is therefore an inadequate summary index. Once again, we examined moderating factors of these effect sizes. Modality (non-aggregated versus deviant-only) across all outcome variables was found to be statistically significant ($Q_B = 15.89$, $p < .05$). The pattern of means was consistent with the hypothesis; the means were 0.24 and 0.51 in the deviant-only group and the non-aggregated group, respectively. We could not conduct further analyses on individual outcome variables as was performed for posttreatment data because of insufficient sample size.

Client Characteristics That Moderate Effect Sizes

Furthermore, we also wanted to examine, within two types of population (preventive versus clinical), if individuals who received non-aggregated treatment had larger mean effect sizes than individuals who received deviant-only group treatment. We predicted that the participants in the non-aggregated treatment format would have larger mean effect sizes compared to participants in the deviant-only group format, only within the preventive population. The preventive group ($n = 27$) included aggressive youth who have not been clinically diagnosed or incarcerated, and the clinical group ($n = 10$) included aggressive youth who have been clinically diagnosed or incarcerated. The overall effect sizes for the preventive and clinical groups are 0.43 and 0.65 respectively. There was heterogeneity in effect sizes of the preventive group ($Q_T = 61.41$, $p < .05$) and therefore, an investigation of factors that might moderate effect size is warranted. As expected, results of the between-group homogeneity statistic indicated that within the less severe, preventive population, the mean for the deviant-only group ($n = 20$, $M = 0.34$) was significantly different from the mean for the non-aggregated group ($n = 7$, $M = 0.67$), $Q_B = 13.45$, $p < .05$. However, no differences in effect size was observed for the more severe clinical population with respect to group format, $Q_B = 0.00044$, ns. The means were 0.59 and 0.77 for deviant-only ($n = 7$) and non-aggregated ($n = 3$) groups respectively. This pattern of results suggest that the less severe, preventive group appears to be more amenable to the effects of treatment format than are the more severe, clinical group.

We also wanted to examine our hypothesis that treatment modality would moderate effect sizes for children (ages 6-12) but not for adolescents (ages 13-18). Because all studies with adolescent clients used deviant-only grouping, we were unable to fully examine this hypothesis. However, for the 24 studies conducted with children, the homogeneity statistic indicated statistically significant heterogeneity in effect sizes, $Q_T = 51.96$, $p < .05$. The mean effect size for the 14 deviant-only group treatment studies ($M = 0.23$) and for the 10 non-aggregated treatment studies ($M = 0.70$) were statistically significantly different, $Q_B = 20.58$, $p < .05$.

Discussion

This study used meta-analytic techniques to summarize treatment outcomes associated with skills training interventions with antisocial youth. Based on both empirical and conceptual arguments, we expected effect sizes would vary systematically based on whether treatment was delivered in the context of aggregated groups of deviant youth versus non-aggregated group or individual treatment. Consistent with our hypothesis, skills training interventions delivered in the context of homogeneous groups of deviant peers produced smaller benefits than did skills training interventions delivered in the context of either individual treatment or mixed groups of prosocial

and deviant peers. This finding held for an analysis of the overall effect size for each study as well as for behavior observations, problem-solving skills, and socially consequential outcomes. Although differences between treatment modality groups in effect sizes for behavior ratings and self-report measures did not reach statistical significance, differences in effect sizes for these outcomes were also in the expected direction.

It is encouraging that the largest effect sizes, irrespective of treatment modality, were found on measures that were socially consequential, defined in terms of outcomes that assess the impact of treatment on important developmental outcomes, such as peer-ratings of acceptance or aggression or recidivism. Although encouraging, these results should be interpreted cautiously based on the few number of studies that included such measures.

Also, consistent with expectations, youth who are "at risk" for serious conditions, such as incarceration in a juvenile facility are more influenced by homogenous grouping than are youth who already experience these conditions. This finding suggests the importance of including prosocial children in interventions for youth at-risk for significant conduct problems. School-based interventions offer the possibility of providing skills training in the context of mixed groups of children with and without problems, as exemplified in the Prinz, Blechman, and Dumas (1994) study. These authors suggested that such mixed groups not only avoid the adverse outcomes associated with deviant peer groups but also engage "high risk children in a supportive, prosocial peer network" (p. 195).

The results of this study must be interpreted in light of certain study limitations. Several potentially important moderator variables of the effectiveness of skills training could not be investigated due to limited information provided in published studies. For example, ethnic and gender differences in responsiveness to skills training interventions and to treatment modality are important to investigate, but too few investigators report results separately by gender and ethnicity to permit a determination of the role of gender and ethnicity in the relationship between treatment responsiveness and treatment modality. Similarly, subtypes of aggressive children, such as proactive and reactive aggressive children, would likely differ in their susceptibility to adverse outcomes associated with grouping deviant peers but could not be examined in this meta-analysis. We also could not determine if treatment modality moderates the effectiveness of skills training with adolescents, due to the absence of intervention studies with this population that utilize non-aggregated modalities.

Approximately half of the studies (51.4%) failed to report follow-up data. If bringing together deviant peers for purposes of skills training results in greater association with deviant peers beyond the duration of the training, the adverse effects of aggregating deviant peers may be greater at follow-up than immediately post-treatment. Dishion et al. (1995) found evidence for adverse effects of aggregating at-risk teens only at post-treatment.

The small number of studies (n = 5) that utilized groups of mixed children required that we combine mixed group-based interventions with individually-provided interventions. These two formats both avoid aggregating deviant peers; however, they introduce a confound in that studies in the differences between aggregated and non-aggregated studies could be a result of differences in group versus individual treatment instead of aggregating deviant peers. However, our decision to combine these two types of non-aggregated skills training studies is supported by the finding that their effect sizes were comparable.

References

* Arbuthnot, J., & Gordon, D. A. (1986). Behavioral and cognitive effects of a moral reasoning development intervention for high-risk behavior-disordered adolescents. Journal of Consulting and Clinical Psychology, 54(2), 208-216.

Arnold, M. E., & Hughes, J. N. (in press). First do no harm: Adverse effects of grouping deviant youth for skills training. Journal of School Psychology.

Beelmann, A., Pfingsten, U., & Losel, F. (1994). Effects of training social competence in children: A meta-analysis of recent evaluation studies. Journal of Clinical Child Psychology, 23(3), 260-271.

Bergin, A. E., & Lambert, M. J. (1978). The evaluation of therapeutic outcomes. In S. L. Garfield & A. E. Bergin (Eds.), Handbook of psychotherapy and behavior change: An empirical analysis (pp. 139-189). New York: Wiley.

* Camp, B. W., Blom, G. E., Hebert, F., Doorninck, W. J. van (1977). "Think Aloud": A program for developing self-control in young aggressive boys. Journal of Abnormal Child Psychology, 5(2), 157-169.

* Dishion, T. J., & Andrews, D. W. (1995). Preventing escalation in problem behaviors with high-risk young adolescents: Immediate and 1-year outcomes. Journal of Consulting and Clinical Psychology, 63(4), 538-548.

* Etscheidt, S. (1991). Reducing aggressive behavior and improving self-control: A cognitive-behavioral training program for behaviorally disordered adolescents. Behavioral Disorders, 16(2), 107-115.

* Feindler, E. L., Ecton, R. B., Kingsley, D., & Dubey, D. R. (1986). Group anger-control training for institutionalized psychiatric male adolescents. Behavior Therapy, 17, 109-123.

* Feindler, E. L., Marriott, S. A., & Iwata, M. (1984). Group anger control training for junior high school delinquents. Cognitive Therapy and Research, 8(3), 299-311.

* Forman, S. G. (1980). A comparison of cognitive training and response cost procedures in modifying aggressive behavior of elementary school children. Behavior Therapy, 11, 594-600.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage Publications.

* Guerra, N. G., & Slaby, R. G. (1990). Cognitive mediators of aggression in adolescent offenders: 2. Intervention. Developmental Psychology, 26(2), 269-277.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. Journal of Educational Statistics, 6, 107-128.

Hedges, L. V., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. S. Hyde & M. C. Linn (Eds.), The psychology of gender: Advances through meta-analysis (pp. 14-50). Baltimore, MD: The Johns Hopkins University Press.

* Hollin, C. R., & Shareen, A. C. (1983). A skills training approach to the reduction of institutional offending. Personality and Individual Differences, 4(3), 257-264.

* Hudley, C. A. (1994). The reduction of childhood aggression using the brainpower program. In J. Michael & D. C. Smith (Eds.), Anger, hostility, and aggression: Assessment, prevention, and intervention strategies for youth (pp. 313-344). Brandon,VT: Clinical Psychology Publishing Company.

* Hudley, C., & Graham, S. (1993). An attributional intervention to reduce peer-directed aggression among African-American boys. Child Development, 64, 124-138.

* Huey, W. C., & Rank, R. C. (1984). Effects of counselor and peer-led group assertive training on Black adolescent aggression. Journal of Counseling Psychology, 31(1), 95-98.

Kazdin, A. E. (1997). A model for developing effective treatments: Progression and interplay of theory, research, and practice. Journal of Clinical Child Psychology, 26, 114-129.

* Kazdin, A. E., Bass, D., Siegel, T., & Thomas, C. (1989). Cognitive-behavioral therapy and relationship therapy in the treatment of children referred for antisocial behavior. Journal of Consulting and Clinical Psychology, 57(4), 522-535.

* Kazdin, A. E., Esveldt-Dawson, K., French, N. H., & Unis, A. S. (1987). Problem-solving skills training and relationship therapy in the treatment of antisocial child behavior. Journal of Consulting and Clinical Psychology, 55(1), 76-85.

* Kazdin, A. E., Siegel, T. C., & Bass, D. (1992). Cognitive problem-solving skills training and parent management training in the treatment of antisocial behavior in children. Journal of Consulting and Clinical Psychology, 60(5), 733-747.

* Kendall, P. C., Reber, M., McLeer, S., Epps, J., & Ronan, K. R. (1990). Cognitive-behavioral treatment of conduct-disordered children. Cognitive Therapy and Research, 14(3), 279-297.

* Kettewell, P. W., & Kausch, D. F. (1983). The generalization of the effects of a cognitive-behavioral treatment program for aggressive children. Journal of Abnormal Child Psychology, 11(1), 101-114.

* Larson, J. D. (1992). Anger and aggression management techniques through the *Think First* curriculum. Journal of Offender Rehabilitation, 18(1/2), 101-117.

* Lee, D. Y., Hallberg, E. T., & Hassard, H. (1979). Effects of assertion training on aggressive behavior of adolescents. Journal of Counseling Psychology, 26(5), 459-461.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. American Psychologist, 48, 1181-1209.

* Lochman, J. E. (1992). Cognitive-behavioral intervention with aggressive boys: Three-year follow-up and preventive effects. Journal of Consulting and Clinical Psychology, 60(3), 426-432.

* Lochman, J. E., & Curry, J. F. (1986). Effects of social problem-solving training and self-instruction training with aggressive boys. Journal of Clinical Child Psychology, 15(2), 159-164.

* Lochman, J. E., & Lampron, L. B. (1986). Cognitive-behavioral interventions for aggressive boys: 7-month follow-up effects. Journal of Child and Adolescent Psychotherapy, 5(1), 15-23.

* Lochman, J. E., Burch, P. R., Curry, J. F., & Lampron, L. B. (1984). Treatment and generalization effects of cognitive-behavioral and goal-setting interventions with aggressive boys. Journal of Consulting and Clinical Psychology, 52(5), 915-916.

* Lochman, J. E., Coie, J. D., Underwood, M. K., & Terry, R. (1993). Effectiveness of a social relations intervention program for aggressive and nonaggressive, rejected children. Journal of Consulting and Clinical Psychology, 61(6), 1053-1058.

* Lochman, J. E., Lampron, L. B., Gemmer, T. C., Harris, S. R., Wyckoff, G. M. (1989). Teacher consultation and cognitive-behavioral interventions with aggressive boys. Psychology in the Schools, 26, 179-188.

* Long, S., & Sherer, M. (1984). Social skills training with juvenile offenders. Child and Family Behavior Therapy, 6(4), 1-11.

* McMichael, P. (1980). School intervention with disadvantaged children. Educational Studies, 6(1), 65-77.

* Milne, J., & Spence, S. H. (1987). Training social perception skills with primary school children: A cautionary note. Behavioural Psychotherapy, 15, 144-157.

* Ollendick, T. H., Hersen, M. (1979). Social skills training for juvenile delinquents. Behaviour Research and Therapy, 17, 547-554.

* Pepler, D. J., Craig, W., & Roberts, W. L. (1995). Social skills training and aggression in the peer group. In J. McCord (Ed.), Coercion and punishment in long-term perspectives (pp. 213-228). New York: Cambridge University Press.

* Prinz, R. J., Blechman, E. A., & Dumas, J. E. (1994). An evaluation of peer coping-skills training for childhood aggression. Journal of Clinical Child Psychology, 23, 193-203.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. American Psychologist, 47, 1173-1181.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. Psychological Methods, 1, 115-129.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). Benefits of psychotherapy. Baltimore, MD: Johns Hopkins University Press.

* Spence, A. J., & Spence, S. H. (1980). Cognitive changes associated with social skills training. Behaviour Research and Therapy, 18, 265-272.

* Spence, S. H., & Marzillier, J. S. (1981). Social skills training with adolescent male offenders: Short-term, long-term and generalized effects. Behaviour Research and Therapy, 19, 349-368.

* Tanner, V. L., & Holliman, W. B. (1988). Effectiveness of assertiveness training in modifying aggressive behaviors of young children. Psychological Reports, 62, 39-46.

* Tremblay, R. E., Masse, L. C., Pagani, L., & Vitaro, F. (1996). From childhood psyical aggression to adolescent maladjustment: The Montreal prevention experiment. In R. D. Peters & R. J. McHahon (Eds.), Preventing childhood disorders, substance abuse, and delinquency. (Banff International Behavioral Sciences Series Vol. 3, pp. 268-298). Thousand Oaks, CA: Sage Publications.

* Vaughn, S. R., Ridley, C. A., Bullock, D. D. (1984). Interpersonal problem-solving skills training with aggressive young children. Journal of Applied Developmental Psychology, 5, 213-223.

* Vitaro, F., & Tremblay, R. E. (1994). Impact of a prevention program on aggressive children's friendships and social adjustment. Journal of Abnormal Child Psychology, 22(4), 457-475.

Vitaro, F., Tremblay, R. E., Kerr, M., Pagani, L., & Bukowski, W. M. (1997). Disruptiveness, friends' characteristics, and delinquency in early adolescence: A test of two competing models of development. Child Development, 68, 676-689.

* Wodarski, J. S., & Pedi, S. J. (1978). The empirical evaluation of the effects of different group treatment strategies against a controlled treatment strategy on behavior exhibited by antisocial children, behaviors of the therapist, and two self-rating scales that measure antisocial behavior. Journal of Clinical Psychology, 34(2), 471-481.

Table 1

Effect Sizes by Treatment Modality

Outcome	Deviant-only		Non-Aggregated		All studies (n = 37)	
	D	n	D	n	D	n
Overall	.42[a]	27	.69[a]	10	.49	37
Behavior ratings	.48	19	.59	10	.52	29
Behavior observations	.39[a]	16	.80[a]	3	.45	19
Self report	.25	16	.42	4	.29	20
Problem-solving skills	.14[a]	11	.66[a]	5	.30	16
Socially consequential measures	.95[a]	10	1.34[a]	3	1.04	13

Note. n in table refers to number of effect sizes for that comparison. Each study can contribute no more than one effect size for each outcome type.

[a] Effect sizes differ based on treatment modality ($p < .05$).